

# Linear Algebra

haruki

October 14, 2024

## Contents

<b>1</b>	<b>Linear Equations</b>	<b>1</b>
1.1	What is Fields?What is Vector Spaces? . . . . .	1
1.2	Systems of Linear Equations . . . . .	3
1.3	Matrices and Elementary Row Operations . . . . .	3
<b>2</b>	<b>Vector Spaces</b>	<b>8</b>
2.1	Subspaces,Basis and Dimension . . . . .	8
2.2	Coordinate and Rank . . . . .	10
<b>3</b>	<b>Linear Transformations and Forms</b>	<b>13</b>
3.1	Linear Transformations and Matrices . . . . .	13
3.2	Linear Functionals and Forms . . . . .	16

## 1 Linear Equations

### 1.1 What is Fields?What is Vector Spaces?

In short,a Field  $F$  should satisfy these properties:

(A) Axioms for addition

(A1) if  $x,y$  in  $F$ ,then  $x + y$  in  $F$

(A2) Addition is commutative:

$$x + y = y + x$$

for all  $x$  and  $y$  in  $F$

(A3) Addition is associative:

$$x + (y + z) = (x + y) + z$$

for all  $x,y$  and  $z$  in  $F$

(A4) There is an element  $0$  in  $F$  such that  $x + 0 = x$ ,for all  $x$  in  $F$

(A5) To every  $x$  in  $F$ , there exists an element  $-x$  in  $F$  such that

$$(-x) + x = 0$$

(M) Axioms for multiplication

(M1) if  $x, y$  in  $F$ , then  $xy$  in  $F$

(M2) Multiplication is commutative:

$$xy = yx$$

for all  $x$  and  $y$  in  $F$

(M3) Multiplication is associative:

$$x(yz) = (xy)z$$

for all  $x, y$  and  $z$  in  $F$

(M4) There is an element  $1 \neq 0$  in  $F$  such that  $1x = x$ , for all  $x$  in  $F$

(M5) To every  $x \neq 0$  in  $F$ , there exists an element  $\frac{1}{x}$  in  $F$  such that

$$\left(\frac{1}{x}\right)x = 1$$

(D) The distributive law

$$x(y + z) = xy + xz$$

for all  $x, y$  and  $z$  in  $F$

Some familiar propositions like cancellation law and the uniqueness of 0 and 1 is omitted here. See the comprehensive proofs in Baby Rudin.

In the following context, we assume the field  $F$  as a subfield of the complex numbers. Otherwise, it could be possible that  $1 + 1 + \cdots = 0$ ! You must not want to see that, which arises from the nonzero characteristic of the field. So we assume the field  $F$  we discuss is a field of characteristic zero.

Our preparation work for vector spaces has done. Let's define it!

**Definition 1.1.** *A vector space consists of the following:*

1. *a field  $F$  for scalars*
2. *a set  $V$  of objects, called vectors*
3. *an operation called vector addition, which satisfies the axioms for additions (A)*
4. *an operation called scalar multiplication, which satisfies:*
  - (a)  *$c\alpha$  is in  $V$  for every  $\alpha$  in  $V$*
  - (b)  *$1\alpha = \alpha$*

- (c)  $(c_1 c_2)\alpha = c_1(c_2)\alpha$   
 (d)  $c(\alpha + \beta) = c\alpha + c\beta$   
 (e)  $(c_1 + c_2)\alpha = c_1\alpha + c_2\alpha$

The scalar multiplication has familiar properties just like the ordinary multiplication. We will dig into this concept more deeply in the following sections.

## 1.2 Systems of Linear Equations

Of Course, the numbers we involved are the elements of a field  $F$ . A system of linear equations is defined as:

$$\begin{array}{ccccccc} A_{11}x_1 + A_{12}x_2 + \dots + A_{1n}x_n & = & y_1 \\ A_{21}x_1 + A_{22}x_2 + \dots + A_{2n}x_n & = & y_2 \\ \vdots & & \vdots & = & \vdots \\ A_{m1}x_1 + A_{m2}x_2 + \dots + A_{mn}x_n & = & y_m \end{array}$$

Obviously, there's no doubt that all the linear combinations of the equations are still linear equations, and every solution of the previous system of linear equations is also a solution of the new one.

However, we want to transform the original system of linear equations to the simplest form, which requires that the solutions should be the same. Therefore, if the original equations are also the linear combinations of the new one, the solutions must be the same, and we've reached our purpose. The two systems of linear equations are called **equivalent**, if each equation in each system is a linear combination of the equations in the other system. And we have the following apparent theorem.

**Theorem 1.1.** *Equivalent systems of linear equations have exactly the same solutions.*

Then all we have to do is to find the suitable way of linear combinations.

## 1.3 Matrices and Elementary Row Operations

To simplify our notations, we introduce a concept:

$$A = \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \dots & A_{nn} \end{bmatrix}$$

We called  $A$  a matrix. Sometimes we use the notation  $\{A_{ij}\}$  to denote it.

By now the matrix is only a pile of numbers, so we endow it some operations:

1. matrix addition:  $(A + B)_{ij} = A_{ij} + B_{ij}$
2. scalar multiplication:  $(cA)_{ij} = cA_{ij}$
3. matrix multiplication:  $(AB)_{ij} = \sum_{r=1}^p A_{ir}B_{rj}$ , where  $AB$  is a  $m \times n$  matrix, and  $A, B$  are  $m \times p, p \times n$  matrices respectively.

The  $n \times n$  identity matrix is denoted by  $I_n$  with entries  $I_{ij}$  that:

$$I_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

which has a significant property:  $IA = AI = A$ .

The first and second operations make the matrices space become a vector space, and the third operation make the matrices space become an algebra, which will simplify our notations greatly.

An important property of matrix multiplication should be pointed out:

**Proposition 1.1.** *The matrix multiplication is associative.*

*Proof.* The proof is direct. If  $A, B, C$  are  $m \times p, p \times q, q \times n$  matrices, then

$$\begin{aligned} ((AB)C)_{ij} &= \sum_{t=1}^q \left( \sum_{r=1}^p A_{ir}B_{rt} \right) C_{tj} = \sum_{t=1}^q \sum_{r=1}^p A_{ir}B_{rt}C_{tj} \\ &= \sum_{r=1}^p A_{ir} \left( \sum_{t=1}^q B_{rt}C_{tj} \right) = (A(BC))_{ij} \end{aligned}$$

□

But it isn't commutative and you can prove it directly as well.

We have possessed the matrix, so let's come back to the system of linear equations. You could find an amazing thing we expect:

$$AX = Y$$

and where

$$A = \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \dots & A_{nn} \end{bmatrix}$$

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

This is just the same thing we discussed above! So the objects concerned will be matrices from now on.

Don't forget the target we said at the end of the subsection 1.2! I will show you that the way we're looking for can be presented as a convenient matrix.

We define three elementary row operations of matrix as following:

1. multiplication of one row of  $A$  by a nonzero scalar  $c$
2. replacement of the  $r$ th row of  $A$  by row  $r$  plus  $c$  times row  $s$ ,  $c$  is any scalar and  $r \neq s$
3. interchange of two rows of  $A$

The elementary matrices come after:

**Definition 1.2.** An elementary matrix  $E_n$  is a matrix which is the consequence of an elementary row operation on the identity matrix  $I_n$ .

A simple but important conclusion reached now:

**Proposition 1.2.** The consequence of an elementary row operation on  $m \times n$  matrix  $A$  is just  $e_m A$ .

Whose proof is direct and I'm lazy to type.

By the definition of elementary row operations, it's apparent that each operation has a **unique** inverse operation to cancel the effect it cause. For a further convenience, we introduce the inverse of a matrix:

**Definition 1.3.** Let  $A$  be a  $n \times n$  matrix. If  $BA = I$ , we call  $B$  a **left inverse** of  $A$ . If  $AC = I$ , we call  $C$  a **right inverse** of  $A$ . If  $BA = AB = I$ , we call  $B$  a **two-side inverse** of  $A$ . You may find that:

$$B = BI = BAC = IC = C$$

So we use  $A^{-1}$  to denote the **inverse** of  $A$ , and  $A$  is called invertible.

It's simple to prove:

**Theorem 1.2.** 1. If  $A$  is invertible, then  $(A^{-1})^{-1} = A$

2. If  $A, B$  are invertible, then  $AB$  is invertible and the inverse is  $B^{-1}A^{-1}$

Then we can formalize the statement above:

**Theorem 1.3.** The elementary matrix is invertible, and its inverse is just the matrix which corresponds to the inverse operation.

Let's come back to the topic again. Why we choose the elementary row operations? Because we have the following theorem:

**Theorem 1.4.** Let  $A^*$  be  $(A, Y)$  for a system of linear equations which we call the **augmented matrix** of the system.  $B^*, A^*$  are called **row-equivalent** if they are connected by a series of elementary row operations, then the two systems are equivalent as a result.

We can write  $A^*, B^*$  in an equation:

$$B^* = E_1 E_2 \dots E_n A^{*1}$$

and an inverse one:

$$A^* = E_n^{-1} E_{n-1}^{-1} \dots E_1^{-1} B^*$$

A more convenient concept comes now:

**Definition 1.4.** An  $m \times n$  echelon matrix  $R$  is called row-reduced if:

- (a) the first nonzero entry in each nonzero row of  $R$  is equal to 1
- (b) each column of  $R$  which contains the leading nonzero entry of some row has all its other entries 0.
- (c) if rows  $1, 2, \dots, r$  are the nonzero rows of  $R$ , and if the leading nonzero entry of row  $i$  occurs in column  $k_i, i = 1, 2, \dots, r$ , then  $k_1 < k_2 < \dots < k_r$

For example,  $I_n$  is just a row-reduced echelon matrix.

We have a theorem below, whose proof is apparent:

**Theorem 1.5.** Every  $m \times n$  matrix over the field  $F$  is row-equivalent to a row-reduced echelon matrix.

And we have done the most of all things of systems of linear equations. You may find that if we use the elementary row operations on the augmented matrix  $A^*$  to make it become row-reduced echelon matrix  $R$ , then we get an explicit solution, because from the definition, we have  $x_{k_1}, x_{k_2}, \dots$ <sup>2</sup> as linear combinations of the rest  $x_m$ , which by arbitrary values of  $x_m$  becomes a solution.

Of course we forgot the situation of no solutions, which can be simply demonstrated by the statement:

**Proposition 1.3.** Let  $R$  be the row-reduced echelon augmented matrix of a system of linear equations, then:

- If there's a first nonzero entry in the last column, there's no solutions
- If not the above condition, the solution is unique if except the last column  $R$  is  $I_n$ , otherwise there are infinite number of solutions.

If you are smart enough, you would find that the uniqueness of solutions means the coefficient matrix  $A$  is an invertible matrix, because:

$$E_1 E_2 \dots E_n A = I_n$$

implies the inverse of  $A$  is:

$$A^{-1} = E_n^{-1} E_{n-1}^{-1} \dots E_1^{-1} I_n$$

---

<sup>1</sup>the foot scripts here don't mean the dimension

<sup>2</sup>the foot scripts label the first nonzero entry of each nonzero row

and the procedure can be taken by hand if we write them together  $(A, I_n)$ , and use the elementary row operations till  $A$  becomes  $I_n$ , then the right is just  $A^{-1}$ .

The last thing about systems of linear equations you need to know is that the solutions of a system of linear equations is a vector space. Let the rest  $x_m$  be a series of values:

$$(1, 0, \dots, 0), (0, 1, \dots, 0), \dots (0, 0, \dots, 1)$$

and a series of solutions come after. You may notice that each solution of the system is a linear combination of the  $n - r$  solutions, which implies the solutions of the system is a vector space that has a dimension of  $n - r$ , which we will detail in the next section.

We finish this section by some properties of matrices:

**Theorem 1.6.** 1. Write  $B$  as  $(\beta_1, \beta_2, \dots, \beta_n)$ , then  $AB = (A\beta_1, A\beta_2, \dots, A\beta_n)$

2.  $A$  is called upper triangle matrix if it only have nonzero entries above the diagonal (contain the diagonal), and the product of two upper triangle matrices is still a upper triangle matrix

3. It will be useful if you notice that  $AE_n$  corresponds to an elementary column operation, which is just a copy version of the row one.



## 2 Vector Spaces

### 2.1 Subspaces, Basis and Dimension

For a further discussion about linear algebra, we will introduce some useful concepts:

**Definition 2.1.** Let  $V$  be a vector space over the field  $F$ .  $W$  is a subset of  $V$ .  $W$  is a subspace of  $V$  if  $W$  is a vector space with vector addition and scalar multiplication on  $V$ .

A simple method to check  $W$  is make an arbitrary linear combination by two arbitrary vectors in  $W$ , which means:

**Theorem 2.1.** If  $\alpha$  and  $\beta$  are vectors in  $W$ , then  $W$  is a subspace if and only if  $c\alpha + \beta$  in  $W$  for each pair of  $\alpha, \beta$  and each scalar  $c$ .

We will introduce the concept “span” in a new way:

**Theorem 2.2.** The intersection of any collections of subspaces is also a subspace.

The proof is easy, so let's skip it.

**Definition 2.2.** Let  $S$  be a set of vectors of  $V$ . The **subspace spanned** by  $S$  is the intersection  $W$  of all subspaces that contains  $S$ . If  $S$  is a finite set, then we call  $W$  the **subspace spanned by the vectors**  $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ .

A more usual form comes from this:

**Theorem 2.3.** The subspace spanned by a non-empty subset  $S$  of a vector space  $V$  is the set of all linear combinations of vectors in  $S$ .

There's an operation of addition defined on the subsets:

**Definition 2.3.** If  $S_1, S_2, \dots, S_k$  is subsets of a vector space  $V$ , the set of all sums

$$\alpha_1, \alpha_2, \dots, \alpha_k$$

of vectors  $\alpha_i$  in  $S_i$  is called the **sum** of the subsets  $S_1, S_2, \dots, S_k$  and is denoted by

$$S_1 + S_2 + \dots + S_k$$

If  $S_1, S_2, \dots, S_k$  are subspaces, then the sum of them is also a subspace.

A further exploration need the concept “linear dependent”:

**Definition 2.4.** A subset  $S$  of a vector space  $V$  is called **linearly dependent** if there exists distinct vectors  $\alpha_1, \alpha_2, \dots, \alpha_n$  in  $S$  and scalars  $c_1, c_2, \dots, c_n$  in  $F$ , not all of which are 0, such that

$$c_1\alpha_1 + c_2\alpha_2 + \dots + c_n\alpha_n = 0$$

A set which is not linear dependent is called **linearly independent**. If  $S$  is only a finite set of vectors, we call these vectors **dependent** or **independent** rather than call  $S$ .



An important remark is that if  $S$  is an infinite set, the concept above didn't mean **any** infinite sum of vectors. We will see this in the following concept.

**Defination 2.5.** A **basis** for  $V$  is a linearly independent set of vectors in  $V$ , which span the  $V$ . The **dimension** of  $V$  is the number of elements in the basis. It's easy to see that the linear combination of each vector is unique

You may notice a theorem is a must to make the definition legitimate.

**Lemma 2.1.** Let  $\alpha_1, \alpha_2, \dots, \alpha_k$  be independent vectors. A vector  $\beta$  is a linear combination of  $\alpha_1, \alpha_2, \dots, \alpha_k$  if and only if  $\beta, \alpha_1, \alpha_2, \dots, \alpha_k$  are dependent vectors.

**Lemma 2.2.** If a nonzero vector  $\beta$  is a linear combination of independent vectors  $\alpha_1, \alpha_2, \dots, \alpha_k$ , there exists  $i$  such that  $\beta, \alpha_1, \alpha_2, \dots, \alpha_{i-1}, \alpha_{i+1}, \dots, \alpha_k$  are independent vectors.

*Proof.* Because  $\beta$  is nonzero, there exists  $i$  such that the coefficient of  $\alpha_i$  is nonzero, which means  $\alpha_i$  is a linear combination of  $\beta, \alpha_1, \alpha_2, \dots, \alpha_{i-1}, \alpha_{i+1}, \dots, \alpha_k$ . If  $\beta, \alpha_1, \alpha_2, \dots, \alpha_{i-1}, \alpha_{i+1}, \dots, \alpha_k$  are dependent, the coefficient of  $\beta$  should be nonzero otherwise  $\alpha_1, \alpha_2, \dots, \alpha_{i-1}, \alpha_{i+1}, \dots, \alpha_k$  are dependent, which means  $\alpha_i$  is a linear combination of  $\alpha_1, \alpha_2, \dots, \alpha_{i-1}, \alpha_{i+1}, \dots, \alpha_k$ , and by the Lemma 2.1, they are dependent.

Therefore the hypothesis is incorrect.  $\square$

**Theorem 2.4.** If  $\beta_1, \beta_2, \dots, \beta_m$  are independent vectors so as  $\alpha_1, \alpha_2, \dots, \alpha_n$ , and each of them is a linear combination of  $\alpha_1, \alpha_2, \dots, \alpha_n$ , then they can replace some of the  $\alpha$  to form a new independent vector set.

*Proof.* We notice that repeat the procedure in Lemma 2.2 we can change the original set by the new set, and by the independence of  $\beta$  we assure the nonzero coefficients happen on  $\alpha$ , then the proof is over.  $\square$

From this theorem, We can prove lots of corollaries.

**Corollary 2.1.** If every element of a finite set of independent vectors is a linear combination of the other, then the number of elements of the former  $m$  is no more than the one of the latter  $n$ .

**Corollary 2.2.** The number of every finite basis is equal.

**Corollary 2.3.** If there exists an infinite basis, then there's no finite basis.

Hence definition of dimension is legitimate now.

A simple example of basis is below. Let  $V$  be  $F^{n^3}$ , then

$$e_i = (0, 0, \dots, 1, \dots, 0) \quad \text{The } i\text{th value is } 1$$

is a basis. It's natural to popularize this to infinite-dimensional space  $F^\infty$ , but it's wrong. Do you remember the remark above? Yes, the “**span**” should be “**finite**”

---

<sup>3</sup>This means  $n$ -triples on field  $F$ , whose operations defined as usual.

as the “**finite**” requirement appears in the definition of linear combination! So it’s impossible to express  $(1, 1, \dots)$  by the method above. To “span” this vector space, we need **orthogonal basis** rather than usual basis, which appears in functional analysis.

How can we form a basis? A useful way to do that is extend a basis of a subspace  $W$ . The feasibility of this method is assured by Theorem 2.4, which says every basis of  $V$  can be replaced partly by a basis of  $W$ . There’s a further theorem:

**Theorem 2.5.** *Let  $W_1, W_2$  be finite-dimensional subspaces of  $V$ , then*

$$\dim W_1 + \dim W_2 = \dim(W_1 \cap W_2) + \dim(W_1 + W_2)$$

*Proof.* We can extend the basis of  $W_1 \cap W_2$  in  $W_1, W_2$  respectively, which forms a basis of  $W_1, W_2$  and can be proved easily.  $\square$

Possessing these theorems, we can learn the topics about “coordinate” and “rank”, which will be explained in the next subsection.

## 2.2 Coordinate and Rank

For convenience, the vector space we discuss below is finite-dimensional. To describe the “coordinate”, we need a basis ordered. Having an ordered basis we can define “coordinate” as following:

**Definition 2.6.** *A vector  $\alpha$  in  $V$  can be uniquely presented as a linear combination of an ordered basis  $\{\alpha_i\}$ , and the coefficient  $x_i$  of each  $\alpha_i$  is called the *ith coordinate* of  $\alpha$  related to the ordered basis  $\{\alpha_i\}$ . We can write the coordinates in a matrix called **coordinate matrix***

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

To do something about coordinate further, we need a theorem about matrix. You may notice we develop the properties of matrix at the time we need it.

**Theorem 2.6.** *Write the square matrix  $A$  as column vectors  $(\alpha_1, \alpha_2, \dots, \alpha_n)$ .  $A$  is invertible if and only if  $\alpha_1, \alpha_2, \dots, \alpha_n$  are linearly independent.*

*Proof.* The proof is easy if you notice the fact that a linear combination of  $\alpha_1, \alpha_2, \dots, \alpha_n$  is just a product of two matrices corresponded to a system of linear equations:

$$c_1\alpha_1 + c_2\alpha_2 + \dots + c_n\alpha_n = A \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

and immediately we finish the proof.  $\square$

Let's come back to the topic. One thing you must notice is the term "relate", which tells us the arbitrariness of the coordinate because of the arbitrary choice of basis. Therefore a natural question is how to change the coordinate when the basis changed.

**Theorem 2.7.** Let  $\{\alpha_i\}, \{\alpha'_i\}$  be two bases.  $P$  is a matrix whose entries is

$$\alpha_i = \sum_{j=1}^n P_{ji} \alpha'_j$$

Then the two coordinates are related by

$$X' = PX$$

or

$$X = P^{-1}X'$$

The invertibility of  $P$  is guaranteed by Theorem 2.6.

Inversely, an invertible matrix  $P$  relate  $\{\alpha_i\}$  to a unique basis  $\{\alpha'_i\}$  that

$$X' = PX$$

and

$$X = P^{-1}X'$$

.

The proof is direct.

The rest properties of coordinate are left to the next section. Therefore let's discuss the term "rank".

**Definition 2.7.** The **row rank** of a matrix  $A$  is the dimension of the vector space<sup>4</sup> spanned by the row vectors of  $A$ . The **column rank** is defined similarly.

**Theorem 2.8.** Row-equivalent matrices have the same row space.

*Proof.* Use Theorem 2.4. □

The similar conclusion about column-equivalent is established as well.

**Theorem 2.9.** The nonzero rows of row-reduced echelon matrix  $R$  is a basis of the row space. So as column-reduced echelon matrix.

If a matrix  $R$  is both a row-reduced echelon matrix and a column-reduced echelon matrix, it's direct to draw the conclusion:

**Theorem 2.10.** The row rank is equal to the column rank. Hence we call them the **rank** of matrix  $A$ .

We can also call the finite dimension of a space spanned by a vector set the **rank** of the vector set.

Why we introduce the rank of a matrix? You will find the answer in the next section.

---

<sup>4</sup>Obviously a subspace of  $F^n$



### 3 Linear Transformations and Forms

#### 3.1 Linear Transformations and Matrices

**Definition 3.1.** Let  $V, W$  be vector spaces over  $F$ . A **linear transformation** from  $V$  to  $W$  is a function  $T$  from  $V$  to  $W$  such that

$$T(c\alpha + \beta) = c(T\alpha) + T\beta$$

for all  $\alpha, \beta$  in  $V$  and  $c$  in  $F$ .

Obviously, the image

**Definition 3.2.** The **null space** of  $T$  is the set of all vectors  $\alpha$  in  $V$  such that  $T\alpha = 0$ .

If  $V$  is finite-dimensional, the **rank** of  $T$  is the dimension of the range of  $T$ , and the **nullity** of  $T$  is the dimension of the null space of  $T$ .

For convenience, we confined ourselves to finite-dimensional vector spaces. We will have an important theorem.

**Theorem 3.1.**  $\text{rank}(T) + \text{nullity}(T) = \dim(V)$

*Proof.* Let  $\{\alpha_i\}$  be an ordered basis of the null space  $W$  with dimension  $m$ . Use theorem 2.4 we can extend it to an ordered basis  $\{\alpha_i\}$  of  $V$ . Now we prove  $T\alpha_i, i = m+1, \dots, n$  is an ordered basis of the range of  $V$ .

$$T\alpha = T(x_1\alpha_1 + x_2\alpha_2 + \dots + x_n\alpha_n) = x_{m+1}T\alpha_{m+1} + \dots + x_nT\alpha_n$$

Hence they span the range.

$$c_1T\alpha_{m+1} + \dots + c_{n-m}T\alpha_n = 0$$

$$T(c_1\alpha_{m+1} + \dots + c_{n-m}\alpha_n) = 0 \rightarrow c_1\alpha_{m+1} + \dots + c_{n-m}\alpha_n \in W$$

$$c_1 = c_2 = \dots = c_{n-m} = 0$$

Hence they are linearly independent. We have finished the proof.  $\square$

We have seen the rank in the previous section. We combine them by the theorem below.

**Theorem 3.2.** A linear transformation is determined uniquely by its action on an ordered basis of  $V$ , which is connected to a unique matrix when an ordered basis of  $W$  is chosen.

*Proof.* Let  $\{\alpha_i\}$  be an ordered basis of  $V$ . Let  $\{\gamma_i\}$  in  $W$  satisfy

$$T\alpha_i = \gamma_i$$

Obviously, for all  $\alpha$  in  $V$ , we have

$$T\alpha = \sum_{i=1}^n x_i\beta_i$$

where  $\{x_i\}$  is the coordinate of  $\alpha$ . This determines  $T$  uniquely. Let  $\{\beta_j\}$  be an ordered basis of  $W$ , we have

$$T\alpha = \sum_{i=1}^n x_i \sum_{j=1}^m A_{ji} \beta_j = \sum_{j=1}^m \beta_j \sum_{i=1}^n A_{ji} x_i$$

Let  $\{y_j\}$  be the coordinate of  $\beta = T\alpha$ , we have

$$y_j = \sum_{i=1}^n A_{ji} x_i \rightarrow Y = AX$$

which implies that  $T$  is connected to the  $m \times n$  matrix  $A$ .  $\square$

Apparently, the rank of matrix is just the rank of the linear transformation, because the nullity of  $T$  is just the solution space of  $AX = 0$  with a connection between coordinate  $X$  and  $\alpha$ .

For a further discussion, we should focus on the algebra of linear transformations.

**Theorem 3.3.** *Let  $V, W, Z$  be vector spaces.  $U, T$  is linear transformations from  $W$  to  $Z$  and from  $V$  to  $W$  respectively. Then the function  $UT$  is a linear transformation from  $V$  to  $Z$ . If  $V$  is invertible, the inverse  $T^{-1}$  is also a linear transformation from  $W$  to  $V$ .*

Because the matrices representation of linear transformations, we can define the addition and scalar multiplication on the linear transformations naturally to make a vector space of linear transformations from  $V$  to  $W$  called  $\mathcal{L}(V, W)$ , whose dimension is  $mn$ . Obviously, the addition, scalar multiplication and function composition (a sense of multiplication) of linear transformation are corresponded to the ones of matrix. The linear transformations from  $V$  to  $V$  is denoted by  $\mathcal{L}(V)$ . We call  $T$  non-singular if  $T\alpha = 0$  implies  $\alpha = 0$ , which implies that the nullity of  $T$  is 0 and  $T$  is one-to-one.

**Theorem 3.4.**  *$T$  is non-singular if and only if  $T$  carries each linearly independent subset of  $W$  onto a linearly independent subset of  $W$ .*

The proof is one-step.

**Theorem 3.5.** *Let  $\dim V = \dim W$ .  $T$  is invertible if and only if  $T$  is non-singular or  $T$  is onto.*

Use Theorem 3.1 and the proof is done.

**Definition 3.3.**  *$V$  is isomorphic to  $W$  if there exists an one-to-one linear transformation  $T$  of  $V$  onto  $W$ , which is called isomorphism.*

Obviously, isomorphism is an equivalence relation on the class of vector spaces.

**Theorem 3.6.**  *$V$  is isomorphic to  $W$  if and only if they have the same dimension.*

*Proof.* Let a linear transformation connect the ordered bases of  $V, W$ , then we have done the proof.  $\square$

This theorem tells that the matrix of an invertible  $T$  is a square matrix.

**Theorem 3.7.** *The linear transformation is invertible if and only if a matrix of it is invertible, and the matrix of it related to the same basis is just the inverse one.*

The proof is direct. It's important to notice that the  $A$  depends on the choice of ordered bases. You may have found that the  $P$  matrix of basis transformation from  $\{\alpha_i\}$  to  $\{\alpha'_i\}$  is just the matrix of a linear transformation  $C$  related to an ordered basis of  $V$  such that

$$\alpha'_i = C\alpha_i = \sum_{j=1}^n P_{ji}^{-1} \alpha_j$$

We want to use  $C$  on  $\alpha$ .

$$C\alpha = \sum_{i=1}^n x_i \alpha'_i = \sum_{j=1}^n \sum_{i=1}^n P_{ji}^{-1} x_i \alpha_j \rightarrow X'' = P^{-1}X$$

If you look it carefully enough, you may find some difference compared to the previous statement. This denotes a different view of basis transformation. The view before is that basis is changed and vectors is unchanged, which is called passive view. The view above is that vectors is changed and basis is unchanged, which is called active view. We will use passive view below.

**Theorem 3.8.** *Let  $\{\alpha_i\}, \{\alpha'_i\}$  be ordered bases of  $V, T \in \mathcal{L}(V)$ .  $A, B$  is the matrices of  $T$  related to them respectively. Then we have*

$$B = P^{-1}AP$$

$$\text{where } X = PX', \alpha'_i = C\alpha_i = \sum_{j=1}^n P_{ji} \alpha_j$$

*Proof.*

$$AX = Y$$

$$P^{-1}APP^{-1}X = P^{-1}Y$$

$$P^{-1}APX' = BX' = Y'$$

$\square$

**Defination 3.4.** *Let  $A, B$  be square matrices.  $B$  is **similar** to  $A$  if there exists an invertible matrix  $P$  such that*

$$B = P^{-1}AP$$

We have

$$C = Q^{-1}BQ = Q^{-1}P^{-1}APQ = (PQ)^{-1}A(PQ)$$

Thus, similarity is an equivalent relation on the set of  $n \times n$  square matrices, which implies that  $B, A$  are corresponded to the same linear transformation related to different ordered bases.

### 3.2 Linear Functionals and Forms

A linear transformation  $f$  from  $V$  to  $F$  is called **linear functional**. We use  $V^*$  to denote  $\mathcal{L}(V, F)$ . Obviously we have

$$\dim V^* = \dim V$$

If  $\{\alpha_i\}$  is an ordered basis of  $V$ , we have a matrix representation of  $f$

$$[a_1 \quad a_2 \quad \dots \quad a_n]$$

where  $a_i = f(\alpha_i)$ .

A natural basis of  $V^*$  comes after, if we define  $f_i(\alpha_j) = \delta_{ij}$ , which is guaranteed by Theorem 3.2. Then we have the theorem below.

**Theorem 3.9.** *If  $V$  is a finite-dimensional vector space and has an ordered basis  $\{\alpha_i\}$ , then there is a unique **dual basis**  $\{f_i\}$  for  $V^*$  such that  $f_i(\alpha_j) = \delta_{ij}$ . For each linear functional  $f$  on  $V$  we have*

$$f = \sum_{i=1}^n f(\alpha_i) f_i$$

and for each vector  $\alpha$  in  $V$  we have

$$\alpha = \sum_{i=1}^n f_i(\alpha) \alpha_i$$

The proof is direct.

**Definition 3.5.** *If  $V$  is a vector space over the field  $F$  and  $S$  is a subset of  $V$ , the **annihilator** of  $S$  is the set  $S^0$  of linear functionals  $f$  on  $V$  such that  $f(\alpha) = 0$  for every  $\alpha$  in  $S$ .*

It's clear that  $S^0$  is a subspace of  $V^*$ .

**Theorem 3.10.** *Let  $V$  be a finite-dimensional vector space and let  $W$  be a subspace of  $V$ , we have*

$$\dim W + \dim W^0 = \dim V$$



*Proof.* Let  $\{\alpha_i\}, i = 1, 2, \dots, k$  be an ordered basis of  $W$  and extend it to an ordered basis of  $V$ . A linear functional  $f$  can be written as a linear combination of the dual basis related to  $\{\alpha_i\}$

$$f = \sum_{i=1}^n f(\alpha_i) f_i$$

For the functionals in  $W^0$  we have

$$\forall \alpha \in W, f(\alpha) = 0 \rightarrow \sum_{i=1}^k c_i f(\alpha_i)$$

Because  $c_i$  is arbitrary, we have

$$f(\alpha_i) = 0, i = 1, 2, \dots, k$$

Therefore  $\{f_i\}, i = k + 1, k + 2, \dots, n$  is an ordered basis of  $W^0$  and we have

$$\dim W + \dim W^0 = \dim V$$

□

In a vector space of dimension  $n$ , a subspace of dimension  $n - 1$  is called a **hyperspace**. Obviously a hyperspace is a null space of a linear functional. Use the theorem above we have

**Corollary 3.1.** *If  $W$  is a  $k$ -dimensional subspace of an  $n$ -dimensional vector space  $V$ , then  $W$  is the intersection of  $n - k$  hyperspaces in  $V$*

*Proof.* From the proof above we learn that  $W$  is the intersection of null spaces of  $f_i, i = k + 1, k + 2, \dots, n$ , then we done the proof. □

**Corollary 3.2.** *If  $W_1, W_2$  are subspaces of finite-dimensional vector space  $V$ , then  $W_1 = W_2$  if and only if  $W_1^0 = W_2^0$ .*

*Proof.* Use Corollary 3.1 □

Another proof of Corollary 3.1 is that consider  $f(\alpha) = 0$  as a system of linear equations, the solution space is  $n - k$ -dimensional with rank of  $k$ .

You may ask if there exists a basis of  $V$  dual to a basis of  $V^*$ . The answer is to consider  $V^{**}$ . Let  $L_\alpha$  be a linear functional on  $V^*$  such that

$$L_\alpha(f) = f(\alpha), \alpha \in V$$

It seems to be a connection to the vector space  $V$ .

**Theorem 3.11.** *Let  $V$  be a finite-dimensional vector space over the field  $F$ . For each vector  $\alpha$  in  $V$  define*

$$L_\alpha = f(\alpha), f \in V^*$$

*The mapping  $\alpha \rightarrow L_\alpha$  is then an isomorphism of  $V$  onto  $V^{**}$*

*Proof.* It's easy to show that the mapping is a linear transformation and non-singular. Use Theorem 3.5.  $\square$

**Corollary 3.3.** *Let  $V$  be a finite-dimensional vector space over the field  $F$ . If  $L$  is a linear functional on the dual space  $V^*$  of  $V$ , then there is a unique vector  $\alpha$  in  $V$  such that*

$$L(f) = f(\alpha)$$

*for every  $f$  in  $V^*$ .*

**Corollary 3.4.** *Let  $V$  be a finite-dimensional vector space over the field  $F$ . Each basis for  $V^*$  is the dual of some basis for  $V$ .*

*Proof.* Each basis for  $V^*$  has a dual basis for  $V^{**}$ . Use Corollary 3.3, there exists a basis for  $V$  such that

$$L_i(f_j) = f_j(\alpha_i) = \delta_{ij}$$

whose dual basis is exactly  $\{f_i\}$ .  $\square$

We usually identify  $\alpha$  with  $L_\alpha$  and say that  $V$  is the dual space of  $V^*$

**Theorem 3.12.** *If  $S$  is a subset of a finite-dimensional vector space  $V$ , then  $(S^0)^0$  is the subspace spanned by  $S$ .*

*Proof.* Let  $W$  be the subspace spanned by  $S$ . Clearly  $W^0 = S^0$ .  $W$  is uniquely determined by  $W^0$  according to Corollary 3.2. You may notice that for  $\alpha$  in  $(S^0)^0$  we have

$$L_\alpha(f) = f(\alpha) = 0$$

which means  $(W^0)^0 = W$ , according to Theorem 3.10.  $\square$

As for infinite-dimensional vector space, we omitted the proof.

It's clear that our definition of hyperspace fails when  $V$  is infinite-dimensional, so we need another definition.

**Definition 3.6.** *If  $V$  is a vector space, a **hyperspace** in  $V$  is a **maximal proper subspace** of  $V$ , which satisfies*

1.  $N$  is a proper subspace of  $V$ .
2. if  $W$  is a subspace of  $V$  which contains  $N$ , then either  $W = N$  or  $W = V$ .

**Theorem 3.13.** *If  $f$  is a non-zero linear functional on the vector space  $V$ , then the null space of  $f$  is a hyperspace of in  $V$ . Conversely, every hyperspace in  $V$  is the null space of a (not unique) nonzero linear functional on  $V$ .*

*Proof.* Clearly  $N_f$  is a proper subspace of  $V$ . Let  $W$  be a subspace of  $V$  which contains  $N_f$  and  $W \neq N_f$ . Choose the vector  $\alpha \in W \setminus N_f$  and a vector  $\beta$  in  $V$  such that  $f(\alpha), f(\beta) \neq 0$ , then we define

$$c = \frac{f(\beta)}{f(\alpha)}$$

let  $\gamma = \beta - c\alpha$ , we have

$$f(\gamma) = f(\beta) - cf(\alpha) = 0$$

which means  $\gamma \in N_f$ . Therefore  $\beta = \gamma + c\alpha \in W, W = V$ .  
Let  $N$  be a hyperspace in  $V$ . Choose a vector  $\alpha \in V \setminus N$ , we have

$$\beta = \gamma + c\alpha, \gamma \in N, \beta \in V$$

The  $\gamma$  and  $\alpha$  is determined uniquely by  $\beta$ . If we have

$$\beta = \gamma' + c'\alpha$$

then

$$(c - c')\alpha = \gamma' - \gamma$$

If  $c - c' \neq 0$ , then  $\alpha \in N$ . Hence  $c = c', \gamma = \gamma'$ . We can define  $g(\beta) = c$ , which is a linear functional on  $V$ , and then  $N$  is the null space of  $g$ .  $\square$

**Lemma 3.1.** *If  $f, g$  are linear functionals on  $V$ , then  $g = cf, c \in F$  if and only if  $N_f \subseteq N_g$ .*

*Proof.* The ‘only if’ part is trivial. If  $f = 0$ , then  $g = 0$ . Thus we assume  $f \neq 0$ . The null space  $N_f$  is a hyperspace of  $V$ . Choose  $\alpha \in V \setminus N_f$  and define

$$c = \frac{g(\alpha)}{f(\alpha)}$$

$$h = g - cf$$

If  $\beta \in V$ , similar to the proof above we have

$$h(\beta) = g(\gamma + c'\alpha) - cf(\gamma + c'\alpha) = c'(g(\alpha) - cf(\alpha)) = 0$$

Hence  $g = cf$   $\square$

**Theorem 3.14.** *Let  $g, f_1, f_2, \dots, f_r$  be linear functionals on a vector space  $V$  with respective null space  $N, N_1, N_2, \dots, N_r$ . Then  $g$  is a linear combination of  $f_1, \dots, f_r$  if and only if  $N_1 \cap N_2 \cap \dots \cap N_r \subseteq N$ .*

*Proof.* The ‘only if’ part is trivial. Suppose the theorem holds for  $r = k - 1$ . If we restrict  $g, f_1, \dots, f_{k-1}$  on  $N_k$  to be  $g', f'_1, \dots, f'_{k-1}$ , it's clear that  $N'_i = N_i \cap N_k, N' = N \cap N_k$ , we have

$$N_1 \cap \dots \cap N_k = N'_1 \cap \dots \cap N'_{k-1} \subseteq N'$$

then use induction hypothesis

$$g' = \sum_{i=1}^{k-1} c_i f'_i$$

Now define  $h$  such that

$$h = g - \sum_{i=1}^{k-1} f_i$$

Then  $h(\alpha) = 0$  for every  $\alpha \in N_k$ . By the proceeding lemma we have  $h = c_k f_k$ , then

$$g = \sum_{i=1}^k f_i$$

□

Before we turn to discuss multilinear forms, let's talk something about the transpose of a linear transformation. It's easy to define the transpose of a matrix

$$A_{ij}^T = A_{ji}$$

and obviously has the properties

1.  $(A^T)^T = A$
2.  $(AB)^T = B^T A^T$

To introduce the transpose of a linear transformation, we need the help of dual space.

**Theorem 3.15.** *Let  $V, W$  be vector spaces over the field  $F$ . For each linear transformation  $T$  from  $V$  into  $W$ , there is a unique linear transformation  $T^T$  from  $W^*$  into  $V^*$  such that*

$$(T^T g)\alpha = g(T\alpha)$$

for every  $g$  in  $W^*$  and  $\alpha$  in  $V$ .

*Proof.*

$$(T^T(cg_1 + g_2))(\alpha) = (cg_1 + g_2)(T\alpha) = cg_1(T\alpha) + g_2(T\alpha)$$

$$T^T(cg_1 + g_2) = cT^T g_1 + T^T g_2$$

The uniqueness is easy to prove. □

We call  $T^T$  be the **transpose** of  $T$ .  $T^T$  is also called the adjoint of  $T$ .

**Theorem 3.16.** *Let  $V, W$  be vector spaces over the field  $F$  and let  $T$  be a linear transformation from  $V$  into  $W$ . The null space of  $T^T$  is the annihilator of the range of  $T$ . If  $V, W$  is finite-dimensional, then*

$$i \text{ rank}(T^T) = \text{rank}(T)$$

ii the range of  $T^T$  is the annihilator of the null space of  $T$ .

*Proof.* Use the definition

$$(T^T g)\alpha = g(T\alpha)$$

It's clear that the null space of  $T^T$  is precisely the annihilator of the range of  $T$ . If  $V, W$  are finite-dimensional, use Theorem 3.1 and Theorem 3.10 we have

$$\text{rank}(T^T) = \dim W^* - \text{nullity}(T^T)$$

$$\text{nullity}(T^T) = \dim W - \text{rank}(T)$$

$$\dim W = \dim W^* \rightarrow \text{rank}(T^T) = \text{rank}(T)$$

If  $\alpha \in N_T$  then we have

$$(T^T g)\alpha = g(T\alpha) = 0$$

implies that the range of  $T^T$  is a subset of the annihilator of  $N_T$ . We have

$$\text{rank}(T^T) = \text{rank}(T) = \dim V - \text{nullity}(T) = \dim N_T^0$$

Thus the range of  $T^T$  is precisely the annihilator of the null space of  $T$ .  $\square$

We know the relation between matrices and linear transformations, thus we have

**Theorem 3.17.** *Let  $V, W$  be vector spaces over the field  $F$  and let  $T$  be a linear transformation from  $V$  into  $W$ . Let  $\{\alpha_i\}, \{\beta_i\}$  be ordered bases for  $V, W$  and let  $\{f_i\}, \{g_i\}$  be the dual bases of them. Let  $A$  be the matrix of  $T$  relative to  $\{\alpha_i\}, \{\beta_i\}$  and let  $B$  be the matrix of  $T^T$  relative to  $\{f_i\}, \{g_i\}$ . We have*

$$B^T = A$$

*Proof.*

$$B_{ji} = (T^T g_i)\alpha_j = g_i(T\alpha_j) = A_{ij}$$

$\square$

Use this we can prove a theorem which we have proven by a different method before.

**Theorem 3.18.** *The row rank is equal to the column rank.*

Let's turn to the topic.

**Definition 3.7.** *A **multilinear function** (tensor) is a function from  $V^r = V \times V \times \cdots \times V$  into  $F$ , which satisfies*

$$L(\alpha_1, \dots, c\alpha_i + \beta_i, \dots, \alpha_r) = cL(\alpha_1, \dots, \alpha_i, \dots, \alpha_r) + L(\alpha_1, \dots, \beta_i, \dots, \alpha_r)$$

*for each  $i$  and every  $\alpha \in V$ . We have the natural definition of addition and scalar multiplication*

$$(L + M)(\alpha_1, \dots, \alpha_r) = L(\alpha_1, \dots, \alpha_r) + M(\alpha_1, \dots, \alpha_r)$$

$$(cL)(\alpha_1, \dots, \alpha_r) = cL(\alpha_1, \dots, \alpha_r)$$

*Then the collection of all multilinear functions on  $V$  is a vector space denoted by  $\mathcal{M}^r(V)$ .*

---

<sup>5</sup>You can define the natural addition and scalar multiplication on  $V^r$  like  $R^n$

We want to construct an ordered basis of  $\mathcal{M}^r(V)$ , we need introduce a concept.

**Definition 3.8.** *The tensor product of  $L, M$  is a multilinear function on  $V^{r+s}$  such that*

$$L \otimes M(\alpha_1, \dots, \alpha_{r+s}) = L(\alpha_1, \dots, \alpha_r)M(\alpha_{r+1}, \dots, \alpha_{r+s})$$

**Theorem 3.19.** *For  $L \in \mathcal{M}^r(V), M \in \mathcal{M}^s(V)$ , we have*

1.  $(cL_1 + L_2) \otimes M = cL_1 \otimes M + L_2 \otimes M$
2.  $L \otimes (cM_1 + M_2) = cL \otimes M_1 + L \otimes M_2$
3.  $L \otimes (M \otimes N) = (L \otimes M) \otimes N$

**Theorem 3.20.** *Let  $V$  be a finite-dimensional vector space and let  $\{\alpha_i\}$  be an ordered basis for  $V$ . Let  $\{f_i\}$  be the dual basis for  $V^*$ . Then*

$$f_{j_1} \otimes f_{j_2} \otimes \dots \otimes f_{j_r}, 1 \leq j_1 \leq n, \dots, 1 \leq j_r \leq n$$

*form a basis for  $\mathcal{M}^r(V)$ .*

*Proof.*

$$L(\alpha) = \sum_{j_1, j_2, \dots, j_r=1}^n A_{1j_1} A_{2j_2} \dots A_{rj_r} L(\alpha_{j_1}, \alpha_{j_2}, \dots, \alpha_{j_r})$$

We shall prove

$$L = \sum_{j_1, j_2, \dots, j_r=1}^n L(\alpha_{j_1}, \alpha_{j_2}, \dots, \alpha_{j_r}) f_{j_1} \otimes f_{j_2} \otimes \dots \otimes f_{j_r}$$

which is apparently correct. The linearly independence is easy to prove.  $\square$

**Definition 3.9.** *Let  $L$  be a multilinear function on  $V^r$ . We say that  $L$  is alternating if  $L(\alpha_1, \dots, \alpha_r) = 0$  whenever  $\alpha_i = \alpha_j, |i - j| = 1$ .*

**Theorem 3.21.** *An alternating multilinear function has the properties:*

1.  $L(\alpha_1, \dots, \alpha_r) = 0$  whenever  $\alpha_i = \alpha_j, i \neq j$
2.  $L(\alpha_1, \dots, \alpha_i, \dots, \alpha_j, \dots, \alpha_r) = -L(\alpha_1, \dots, \alpha_j, \dots, \alpha_i, \dots, \alpha_r)$

*Proof.* Let  $i < j$ . If  $j = i + 1$ , we have

$$\begin{aligned} L(\alpha_1, \dots, \alpha_i + \alpha_j, \alpha_i + \alpha_j, \dots, \alpha_r) &= L(\alpha_1, \dots, \alpha_i, \alpha_i, \dots, \alpha_r) + L(\alpha_1, \dots, \alpha_i, \alpha_j, \dots, \alpha_r) \\ &\quad + L(\alpha_1, \dots, \alpha_j, \alpha_i, \dots, \alpha_r) + L(\alpha_1, \dots, \alpha_j, \alpha_j, \dots, \alpha_r) = 0 \end{aligned}$$

$$L(\alpha_1, \dots, \alpha_i, \alpha_j, \dots, \alpha_r) = -L(\alpha_1, \dots, \alpha_j, \alpha_i, \dots, \alpha_r)$$

If  $j > i + 1$ , we need commute one-by-one. Suppose  $k = j - i$ , then we need commute  $2k - 1$  times. Thus we have

$$\begin{aligned} L(\alpha_1, \dots, \alpha_i, \dots, \alpha_j, \dots, \alpha_r) &= (-1)^{2k-1} L(\alpha_1, \dots, \alpha_j, \dots, \alpha_i, \dots, \alpha_r) \\ &= -L(\alpha_1, \dots, \alpha_j, \dots, \alpha_i, \dots, \alpha_r) \end{aligned}$$

If  $\alpha_i = \alpha_j$ ,  $i \neq j$ , we commute to make them adjacent, then we have

$$L(A) = -L(B) = 0$$

□

**Defination 3.10.** For an ordered  $n$ -tuple  $a_1 a_2 \cdots a_n$ , each pair of number forms an ordered pair. Let  $j = i + 1$ , if  $a_i > a_j$ , then we call this pair an **inversion**. The number of all inversion is called the **inversion number** of the  $n$ -tuple.

**Definition 3.11.** A permutation is called an **odd permutation** if its inversion number is odd, and is called an **even permutation** if its inversion number is even.

**Theorem 3.22.** Transposition change the parity of a permutation.

The proof is easy so we omitted it.

**Theorem 3.23.** A permutation is a product of transpositions.

*Proof.* We can use transposition to change a permutation to  $123 \cdots n$ , and then the inverse is also a product of transpositions. □

Thus, The parity of the number of transpositions you need to change a permutation to  $123 \cdots n$  is unique, because the parity of a permutation is unique. Therefore we can define the signal of permutations.

**Definition 3.12.** The signal of a permutation is 1, if it's an even permutation, and is  $-1$ , if it's an odd permutation.

This also tells us that  $L(\alpha_{\sigma(1)}, \alpha_{\sigma(2)}, \dots, \alpha_{\sigma(r)}) = \text{sgn}(\sigma) L(\alpha_1, \dots, \alpha_r)$ , and the definition about alternating is legitimate because of the uniqueness of the parity of the number of transpositions the permutation has.

**Defination 3.13.** A multilinear function  $L$  on  $V^r$  is called a linear  $r$ -form if it is alternating.

The collection of all  $r$ -forms is obviously a subspace of  $\mathcal{M}^r(V)$ , we denote it by  $\bigwedge^r(V)$ .

**Theorem 3.24.** If  $L$  is a non-zero  $r$ -form, then  $r \leq n$ . The dimension of  $\bigwedge^r(V)$ ,  $1 \leq r \leq n$  is  $\binom{n}{r}$ .

*Proof.* There exist  $\binom{n}{r}$   $r$ -shuffle, which is a  $r$ -tuple  $j_1 j_2 \cdots j_r$  with  $j_1 < j_2 < \cdots < j_r$ . For the coordinate of  $f_{j_{\sigma(1)}} \otimes \cdots \otimes f_{j_{\sigma(r)}}$ <sup>6</sup>, we have

$$L(\alpha_{j_{\sigma(1)}}, \dots, \alpha_{j_{\sigma(r)}}) = \text{sgn}(\sigma) L(\alpha_{j_1}, \dots, \alpha_{j_r})$$

and therefore we have

$$L = \sum_{\text{shuffles}-J} L(\alpha_{j_1}, \dots, \alpha_{j_r}) D_J$$

$$D_J = \sum_{\sigma} f_{j_{\sigma(1)}} \otimes \cdots \otimes f_{j_{\sigma(r)}}$$

$D_J$  are obviously linearly independent, hence forms a basis of  $\bigwedge^r(V)$ . We have

$$L = \sum_J c_J D_J$$

□

**Definition 3.14.**  $\pi_r$  is a linear transformation from  $\mathcal{M}^r(V)$  into  $\bigwedge^r(V)$  such that

$$\pi_r(L) = \sum_{\sigma} \text{sgn}(\sigma) L_{\sigma}$$

$$L_{\sigma}(\alpha_1, \dots, \alpha_r) = L(\alpha_{\sigma(1)}, \dots, \alpha_{\sigma(r)})$$

If  $L \in \bigwedge^r(V)$ ,  $\pi_r(L) = r!L$

We can write  $D_J$  as  $\pi_r(f_{j_1} \otimes \cdots \otimes f_{j_r})$ <sup>7</sup>. The proof is easy.

**Corollary 3.5.** If  $V$  is a  $n$ -dimensional vector space, then  $\bigwedge^n(V)$  is a vector space with dimension 1. If  $T$  is a linear transformation on  $V$ , then there exists a unique  $c$  such that

$$L(T\alpha_1, \dots, T\alpha_n) = cL(\alpha_1, \dots, \alpha_n)$$

for all  $L \in \bigwedge^n(V)$

*Proof.* Apparently, define  $L_T(\alpha_1, \dots, \alpha_n) = L(T\alpha_1, \dots, T\alpha_n)$  and  $L_T$  is a  $n$ -form. For a basis  $M$ , we have  $L = aM$  and  $M_T = cM$ , then

$$L_T = (aM)_T = aM_T = c(aM) = cL$$

You may notice that  $D_J$  here is just  $\pi_n(f_1 \otimes \cdots \otimes f_n)$ , and we have

$$c = \sum_{\sigma} \text{sgn}(\sigma) A_{1\sigma(1)} \cdots A_{n\sigma(n)}$$

where  $A$  is the matrix of  $T$  related to the basis  $\{\alpha_i\}$ . We call  $c$  the **determinant** of the transformation  $T$  and denote it as  $\det T$ . It's clear that the matrix determinant comes after, and is independent of the basis. □

<sup>6</sup>We only consider permutations because the alternating property

<sup>7</sup>You may notice that the existence of alternating multilinear functions is proven when we construct the  $D_J$



**Theorem 3.25.** *The determinant has some important properties:*

1.  $\det AB = \det A \det B$
2. *The determinant of upper-triangle matrix is the product of diagonal elements  $a_{11} \cdots a_{nn}$*
3.  $\det A^T = \det A$

*Proof.* The last two statements are easy. We prove the first.

$$\begin{aligned} \det AB &= \sum_{\sigma} \operatorname{sgn}(\sigma) \left( \sum_{k_1=1}^n a_{1k_1} b_{k_1\sigma(1)} \right) \cdots \left( \sum_{k_n=1}^n a_{nk_n} b_{k_n\sigma(n)} \right) \\ &= \sum_{\sigma} \sum_{k_1=1}^n \sum_{k_n=1}^n \operatorname{sgn}(\sigma) a_{1k_1} \cdots a_{nk_n} b_{k_1\sigma(1)} \cdots b_{k_n\sigma(n)} \end{aligned}$$

Every term has  $k_i = k_j$  cancels because of  $\operatorname{sgn}(\sigma)$ , thus  $k_1, k_2, \dots, k_n = \tau(1), \dots, \tau(n)$ .

$$\begin{aligned} \det AB &= \sum_{\sigma} \sum_{k_1=1}^n \sum_{k_n=1}^n \operatorname{sgn}(\sigma) a_{1k_1} \cdots a_{nk_n} b_{k_1\sigma(1)} \cdots b_{k_n\sigma(n)} \\ &= \sum_{\sigma} \sum_{\tau} \operatorname{sgn}(\sigma) a_{1\tau(1)} \cdots a_{n\tau(n)} b_{\sigma^{-1}\tau(1)1} \cdots b_{\sigma^{-1}\tau(n)n} \\ &= \sum_{\sigma} \sum_{\tau} \operatorname{sgn}(\sigma^{-1}\tau) \operatorname{sgn}(\tau) a_{1\tau(1)} \cdots a_{n\tau(n)} b_{\sigma^{-1}\tau(1)1} \cdots b_{\sigma^{-1}\tau(n)n} \\ &= \det A \det B \end{aligned}$$

□